

```
import pandas as pd
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import PorterStemmer
from nltk.stem import WordNetLemmatizer
import string

nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')

csv_file_path =
'/content/fifa_world_cup_2022_tweets.csv'
df = pd.read_csv(csv_file_path)

def clean_text(text):
    words = word_tokenize(text)

    words = [word.lower() for word in words]

    words = [word for word in words if word not in
string.punctuation]

    stop_words = set(stopwords.words('english'))
    words = [word for word in words if word not in
stop_words]

    lemmatizer = WordNetLemmatizer()
    words = [lemmatizer.lemmatize(word) for word
in words]

    cleaned_text = ' '.join(words)

    return cleaned_text
```

```
df['cleaned_tweet'] =  
df['Tweet'].apply(clean_text)  
  
cleaned_csv_file_path =  
'/content/fifa_world_cup_2022_tweets1.csv'  
df.to_csv(cleaned_csv_file_path, index=False)
```

```
import pandas as pd  
from sklearn.feature_extraction.text import  
CountVectorizer  
from sklearn.model_selection import  
train_test_split  
from sklearn.svm import SVC  
from sklearn.metrics import accuracy_score,  
classification_report  
  
csv_file_path =  
'/content/fifa_world_cup_2022_tweets1.csv'  
df = pd.read_csv(csv_file_path)  
  
X = df['cleaned_tweet']  
y = df['Sentiment']  
  
vectorizer = CountVectorizer()  
X_bow = vectorizer.fit_transform(X)  
  
X_train, X_test, y_train, y_test =  
train_test_split(X_bow, y, test_size=0.2,  
random_state=42)  
  
svm_classifier = SVC(kernel='linear')  
svm_classifier.fit(X_train, y_train)  
  
y_pred = svm_classifier.predict(X_test)
```

```
accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test,
y_pred)

print(f"Accuracy: {accuracy}")
print("Classification Report:")
print(classification_rep)
```

Accuracy: 0.7036625971143174

Classification Report:

	precision	recall	f1-score	support
negative	0.73	0.69	0.71	1149
neutral	0.65	0.68	0.66	1648
positive	0.74	0.74	0.74	1708
accuracy			0.70	4505
macro avg	0.71	0.70	0.70	4505
weighted avg	0.71	0.70	0.70	4505